SimianTools: A freely available software suite for high throughput glycomics mass spectrometry data

René Ranzinger, Brent Weatherly, Khalifeh AlJadda, Ki Tae Myoung, Mindy Porterfield, Michael Tiemeyer and William S. York

Complex Carbohydrate Research Center, University of Georgia, Athens, GA, USA

Introduction

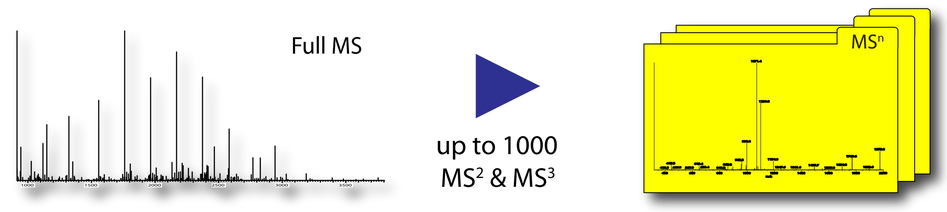
Over the last decades the approaches for analyzing and identifying glycan structures using mass spectrometry have evolved from the study of single purified glycans towards the analysis of complex mixtures. However, most interpretation of the resulting datasets is still done manually. Therefore, comprehensive analysis of complex biological samples, producing several thousand mass spectra, consumes significant amount of time which consequently reduces the number of samples that can be analyzed. Therefore, development of high-throughput glycomics analysis software is crucial to efficiently tackle the large amount of data produced.

Samples

Glycans are released from their protein backbone, either enzymatically or chemically, and permethylated for enhanced sensitivity and structural determination. The mixture of free glycans is dissolved in 1 mM NaOH / 50 % methanol and infused directly into a linear ion trap mass spectrometer (Thermo LTQ Orbitrap Discovery) using a nanoelectrospray source. Many isobaric glycans with different structural features can co-exist at a given m/z. Unique fragment ions from MS^n data can distinguish the different isoforms. Therefore MS^n data is essential for detailed structure elucidation.

Mass spectrometry data

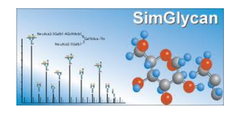
Glycan intact parent masses are determined by high resolution Full MS in the Orbitrap. Structural determinations are made by fragmentation of peaks 3 fold over baseline using collision induced dissociation (CID) in the ion trap.



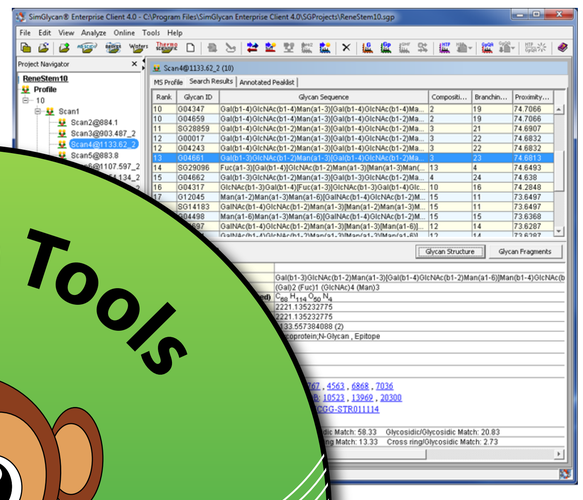
Analysis option 1: SimianTools annotation

- Import MS data files (e.g. mzXML)
- Integrated structual assignment based on fragmentation finger-print matching
- Algorithm is based on GlycoWorkbench annotation (http://code.google.com/p/glycoworkbench/)
- Extention of the algorithm for new features (e.g. neutral loss, under-methylation)
- Use curated glycan structures from GlycO ontology or from customized databases based on GlycomeDB for annotation

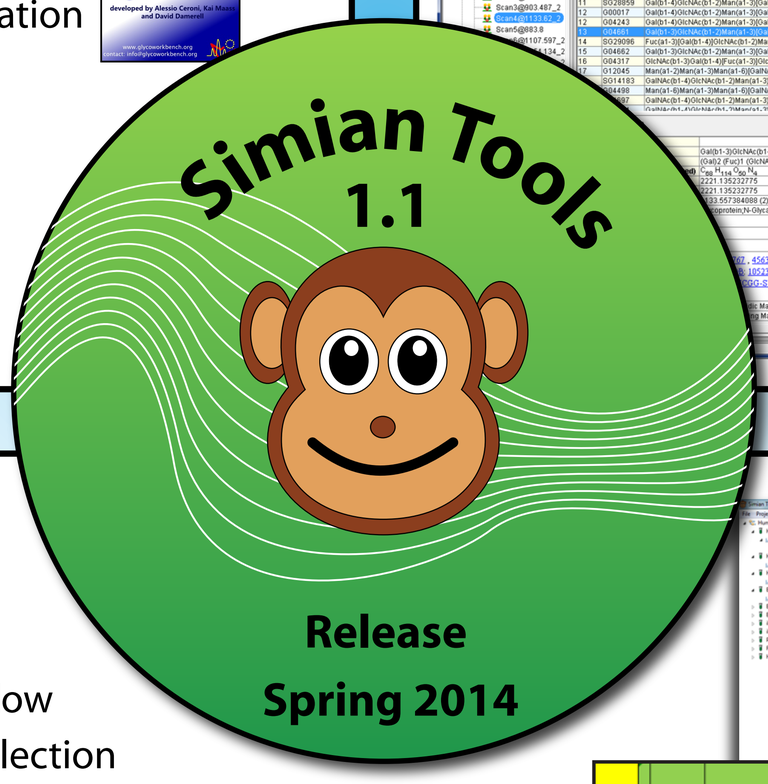
Analysis option 2: SimGlycan™



http://www.premierbiosoft.com/glycan/

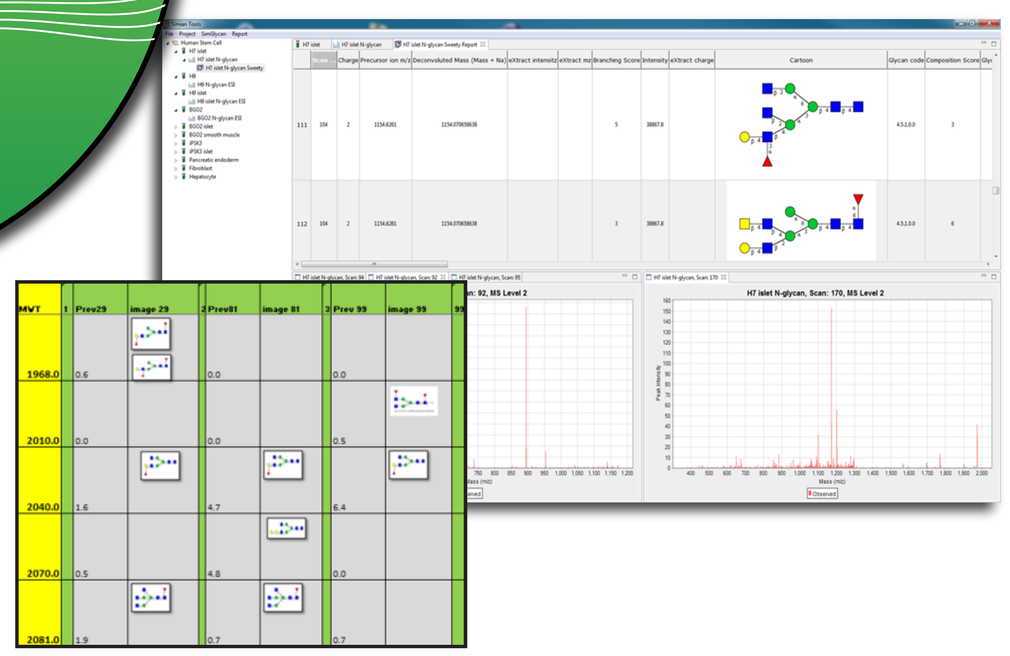


Commercial software application for the annotation of glycomics MS data. Similar to Sequest or Mascot, spectra are annotated with structures from a database. The program allows exploring results and exporting them to CSV files. These files can then be imported to SimianTools.



SimianTools

- Stand-alone JAVA program
- Imports SimGlycan™ annotation data
- Enriches annotation with CFG cartoons that allow easy visual inspection and user defined (de-)selection of the structural assignments
- Enhances annotation information with key numbers and classifiers derived from the spectra or structure
- Integrated spectrum viewer for MS profile and MS^n spectra
- Side by side comparison of datasets from multiple samples or experiments
- Allows organization and storage of experimental data and meta information



http://glycomics.ccrc.uga.edu/simiantools/

Next Steps

- Public release of version 1.1 in Spring 2014
- Full MS spectrum viewer with structure annotations
- MS^n spectrum viewer with fragment annotations
- Support of annotation of LC-MS data
- Extend application for identification of GAG MS data
- Capture more meta information to make the data set MIRAGE (**M**inimum **I**nformation **R**equired for **A** **G**lycomics **E**xperiment) compliant
- Integration of quantification strategies and algorithms
- Import and export to GlycoWorkbench

The University of Georgia

Complex Carbohydrate Research Center

National Institutes of Health